Algorithm to Facilitate Inter-Instrument Comparison of Raman Spectral Libraries

NIST is developing instrument-corrected Raman spectral libraries for homeland defense applications. Raman spectroscopy is an attractive analytical technique for the identification of unknown chemical compounds because it requires no sample preparation, can be used through common, transparent sample containers, and, with current technological innovations, is very portable.

S. Choquette (Div. 839) and A. Kearsley (ITL)

Raman spectroscopy is an emission process, and each instrument design has a unique response function. As a result, Raman libraries of chemical compounds are currently vendor-specific and not transportable between systems. Since 2000, NIST has been involved with the development of instrument performance standards (Standard Reference Materials [SRMs]) to correct for instrument response and has recently been funded by the Department of Homeland Security (DHS) to develop an instrument-corrected Raman spectral database. Because the Raman spectrum of a neat compound is an intrinsic property of the material, this primary data activity should lead to a standardized, universally useable Raman database. This approach is both time consuming and costly as the purity of each compound must be assessed before its spectrum is included in the library. To address the current need for Raman libraries, we are researching algorithmic methods that will enable the use of an existing instrument-specific library with multiple vendor systems. This work is in collaboration with the Federal Bureau of Investigations (FBI) Counterterrorism and Forensic Science unit at Quantico, VA.

The algorithm is based upon a peak-selection method developed at NIST for analysis of standard polymer samples using matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS).

A sample spectrum is measured and then processed into a series of peaks and troughs. The most statistically significant peaks of the sample are aligned with the most closely matched peaks in the library. Orthogonal distance regression is used to determine whether or not the blended data produces a match of the unknown to any compound in the library. The technique does not require that the unknown have the same data spacing, resolution, or even calibration as the library spectra.

The algorithm was tested using a sample set provided by our FBI collaborators. These samples consisted of 34 white powders that might be expected to be found in any household, e.g., talc, baking powder, or powdered milk. Several of these materials do not have Raman spectra (e.g., NaCl) and many fluoresce. The samples were measured on two commercial, portable Raman systems with varying resolution, spectral coverage, and sensitivity.

This NIST-developed algorithm had a 92 % success rate for matching the spectrum of compound run on system 1 with a spectrum from a library generated on system 2 (and vice-versa). By comparison, traditional library search routines had less than 15 % success rate in this exercise.

Impact: Since the presentation of this work at FACSS 2005 (October), we have been approached by three major vendors of Raman systems for information regarding this method.

Future plans:
We plan to create a web-based version that will enable users to compare their spectra with NIST's Raman database, which is under development. We also plan to develop PC-based versions of the algorithm that can be used in conjunction with commercial spectroscopy software packages.